

IMPROVED EFFICIENCY ASSESSMENT IN NETWORK DEA THROUGH INTERVAL DATA ANALYSIS: AN EMPIRICAL STUDY IN AGRICULTURE

NAN ZHANG¹, AMIR KALHOR², ROZA AZIZI³ AND REZA KAZEMI-MATIN^{2,*}

Abstract. Conventional Network Data Envelopment Analysis (NDEA) models often make an assumption of data precision, which frequently does not align with the realities of many real-world scenarios. When dealing with ambiguous data, whether it involves input, output, or intermediate products represented as bounded or ordinal data, the accurate assessment of efficiency scores poses a significant challenge. This study addresses the crucial issue of handling interval data within NDEA structures by introducing an innovative methodology that integrates both optimistic and pessimistic strategies. Our proposed methodology goes beyond the mere determination of upper and lower bounds for efficiency scores; it also incorporates target-setting and improvement approaches. Through the calculation of interval efficiency for each decision-making unit (DMU), our approach offers a comprehensive framework for efficiency classification. To underscore the effectiveness of this methodology, the study presents empirical evidence through a case study in the agriculture industry. The results not only showcase the advantages of our proposed methodology but also emphasize its potential for practical application in diverse and complex real-world contexts.

Mathematics Subject Classification. 90B30.

Received April 7, 2023. Accepted September 19, 2023.

1. INTRODUCTION

Efficiency analysis is a widely employed technique in the performance evaluation domain that aims to assess the relative performance of decision-making units (DMUs) across diverse industries. One prominent non-parametric method used for measuring the relative efficiency of DMUs by comparing their input-output relationships is Data Envelopment Analysis (DEA), which was introduced by Charnes *et al.* [1]. DEA has found applications in diverse sectors, including healthcare, education, finance, agriculture and manufacturing. DEA models are typically based on the assumption that input and output data are precise and deterministic. However, in practical scenarios, input and output data may be imprecise or uncertain due to measurement errors, rounding errors, or subjective judgments. Interval data analysis has emerged as a promising technique to address this uncertainty by incorporating interval-valued data into DEA models.

Keywords. Data Envelopment Analysis, network production systems, interval data, efficiency measurement.

¹ Xingzhi College of Xi'an University of Finance and Economics, Xi'an, Shaanxi, P.R. China.

² Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran.

³ Ivey Business School, Western University, London, Ontario, Canada.

*Corresponding author: rkmartin@kiauo.ac.ir

The presence of significant variation ranges in datasets often leads to an increase in error levels during efficiency evaluation. Furthermore, data can also be uncertain, missing, or sequential, among other possibilities. To address these scenarios, an evaluation process can be employed that considers data within an interval. Interval data analysis deals with uncertain input and output data by representing them as intervals instead of exact values. By adopting this approach, interval data analysis provides a more realistic representation of the true values of input and output data and generates more accurate results in DEA models. Interval data analysis has been widely studied for the last twenty years. Cooper *et al.* [2] pioneered the concept of imprecise data envelopment analysis (IDEA). Imprecise data refers to a combination of interval and ordinal data. Traditional DEA models become nonlinear when imprecise data is present. To deal with this issue, evaluation procedures utilize upper and lower boundaries for each imprecise input and output, and this approach results in one upper and one lower boundary for the efficiency score of each DMU.

A literature review reveals that most IDEA studies introduce models and employ transformation and variable alteration techniques to convert nonlinear models into linear ones. Table 1 summarizes the progress of the most influential IDEA models from Cooper *et al.* [2] to the present day.

Traditional DEA models, including those mentioned earlier, treat DMUs as “black boxes” where only input and output values are used to assess their performance, and the internal operations and structure of the DMUs do not affect the performance evaluation. However, Färe and Grosskopf [42] introduced a new perspective that considers the impact of a DMU’s internal operations and structure on performance evaluation in DEA. This approach is called network data envelopment analysis (NDEA) and has been widely used by researchers in various fields over the past two decades. Kao [43] compiled and categorized almost all publications related to this topic. Kao [44] provides a comprehensive explanation of NDEA, including its model development and practical applications.

NDEAs can be classified into three modes based on the internal structure of the DMUs being evaluated: parallel, series, or mixed (general). In the parallel mode, intermediate products are not exchanged between processes, and each process operates independently. Färe *et al.* [45] were among the early researchers in this field and developed the parallel mode. They used a nonparametric approach to assess the efficiency of 57 farms in Southern Illinois that were producing four types of crops simultaneously in 1994. Cook *et al.* [46] introduced a model to evaluate systems where multiple functions are performed, and Cook and Hababou [47] studied the performance of banks whose primary functions were sales and services. Jahanshahloo *et al.* [48] evaluated the performance of 39 branches of Iranian commercial banks that shared inputs and performed functions related to deposit, sales, and services. Yu and Fan [49] studied the performance of 24 bus companies that provided different urban and highway services. Kao [50] developed a parallel DEA model to measure both DMU efficiency and their corresponding process efficiency. Kao [51] later used this approach to evaluate the efficiency of 52 chemistry departments in England. Kremantzis *et al.* [52] discussed the incorporation of hierarchical network structures in a multi-function parallel system.

The series state is characterized by processes that can take intermediate products or exogenous input as input to produce final output or other intermediate products. Truett *et al.* [53] presented a value-based model that exemplifies a serial process where multiple outputs from a process can be the input of the next process. Amirteimoori and Kordrostami [54] presented aggregation efficiency models that assess the performance of a series system, where the objective function is defined as the weighted mean of the efficiencies of p processes. Park and Park [55] suggested a method to evaluate aggregative efficiency in multi-period production systems. Tsutsui and Goto [56] used a weighted slack-based measure (SBM) to examine the performance of electric power companies in the US. Kao [57] proposed a relational model to measure the efficiency of DMUs and their processes simultaneously. Wei and Chang [58] introduced the Optimal System Design (OSD) Data Envelopment Analysis (DEA) model to optimally design series-network systems of DMUs in terms of profit maximization given the DMUs’ total budget. Fang [59] discovered an error in the Wei and Chang model and suggested a new approach to derive the DMUs’ optimal budgets and to verify the existence of budget congestion for the OSD network DEA models. Lin and Liu [60] extended the multiplier series DEA model to work with the directional distance function, allowing it to handle negative data and be non-oriented.

TABLE 1. A review of studies on imprecise data envelopment analysis.

Publication date	Author(s)	Main contribution to the literature
1999	Cooper <i>et al.</i> [2]	Introduces Imprecise DEA, a novel approach that combines interval and ordinal data, and demonstrates the transformation of nonlinear models into linear ones using appropriate mathematical techniques.
1999	Kim <i>et al.</i> [3]	Converts nonlinear problems into linear ones through scale transformation and variable alteration.
2001	Cooper <i>et al.</i> [4]	Transforms nonlinear models into linear ones without requiring access to maximum data values, contrasting with Cooper <i>et al.</i> [2].
2002	Despotis and Smirlis [5]	Introduces a novel transformation method, utilizing convex combinations of interval endpoints and variable alterations to convert nonlinear models into linear ones. Also introduces upper and lower bounds of efficiency.
2002	Entani <i>et al.</i> [6]	Offers both an optimistic (best-case) and pessimistic (worst-case) perspective for evaluating efficiency values when dealing with interval data.
2002	Lee <i>et al.</i> [7]	Introduces an additive model for imprecise data and converts it into a linear model through transformative techniques.
2003	Zhu [8]	Introduces a method to reduce computational burden when converting nonlinear IDEA models to linear ones, building on prior approaches.
2004	Park [9]	Introduces a simplified method for converting nonlinear models into linear models solely through variable alternation.
2004	Zhu [10]	Demonstrates the straightforward transformation of interval and ordinal data into exact data, enabling the application of conventional DEA models to analyze them.
2004	Jahanshahloo <i>et al.</i> [11]	Suggests an adapted model, the FDH model for interval data, and proposes upper and lower bounds for efficiency levels within this model.
2004	Jahanshahloo <i>et al.</i> [12]	Evaluates the return to scale for DMUs with interval data, specifically those that are always efficient.
2004	Jahanshahloo <i>et al.</i> [13]	Calculates the radius of stability for all DMUs with interval data.
2005	Amirteimoori and Kordrostami [14]	Introduces a model for assessing the efficiency of multi-component structures using imprecise data.
2005	Wang <i>et al.</i> [15]	Introduces models for upper and lower bounds of efficiency with interval data, emphasizing that different constraint sets lead to distinct efficiency frontiers, limiting direct DMU efficiency score comparisons.
2005	Haghighat and Khorram [16]	In the presence of interval data in DEA, the minimum and maximum number of efficient units are determined by the most pessimistic (maximum) and optimistic (minimum) values within the interval data, respectively.
2006	Kao [17]	Suggests two-level mathematical programming models for computing upper and lower bounds of efficiency when dealing with imprecise data.
2006	Smirlis <i>et al.</i> [18]	Utilizes an interval DEA-based method to handle situations with missing data. This approach involves substituting intervals for missing data and subsequently calculating interval efficiency scores for DMUs with incomplete information.
2006	Farzipoor Saen [19]	Proposes procedures for selecting the optimal technology when dealing with a mix of ordinal and cardinal data.
2006	Farzipoor Saen [20]	Offers procedures for ranking the best technology when both ordinal and cardinal data are present.

TABLE 1. continued.

Publication date	Author(s)	Main contribution to the literature
2007	Farzipoor Saen [21]	Proposes procedures for selecting the best supplier when dealing with a combination of ordinal and cardinal data.
2007	Kazemi Matin <i>et al.</i> [22]	Introduces an additive model for evaluating the technical efficiency of DMUs in the context of interval data. Utilizes the translation invariant property of the additive model to transform the nonlinear model into a linear one.
2007	Park [23]	Examines imprecise data in DEA and presents two distinct programs for computing the upper and lower bounds of efficiency.
2009	Jahanshahloo <i>et al.</i> [24]	Introduces a generalized model for interval data utilizing a domination structure.
2009	Farzipoor Saen [25]	Introduces a novel pair of DEA models for selecting the best suppliers when dealing with imprecise data and nondiscretionary factors concurrently.
2010	Park [26]	Analyzes the relationship between envelopment and multiplier forms of DEA models and their respective solutions in the context of Interval Data Envelopment Analysis (IDEA).
2011	Azizi and Ganjeh-Ajirlu [27]	Suggests a pair of DEA models for assessing relative efficiency in scenarios involving nondiscretionary and imprecise data.
2011	Farzipoor Saen [28]	Develops modifications to traditional DEA models to address media selection problems when dealing with both imprecise data and flexible factors.
2011	Farzipoor Saen [29]	Introduces a model for selecting optimal markets that simultaneously accounts for weight restrictions, dual-role factors, and imprecise data.
2011	Kao and Liu [30]	Presents an approach for evaluating interval efficiency within two-stage network systems when dealing with fuzzy data.
2012	Emrouznejad <i>et al.</i> [31]	Introduces nonparametric multiplicative and general models designed to handle interval inputs and outputs.
2013	Zhu and Zhou [32]	Examines the impact of imprecise data within a two-stage network structure.
2013	Bouzembrak <i>et al.</i> [33]	Extends a comprehensive possibilistic linear model for supply chain systems with imprecise inputs and presents a solution approach for a possibilistic mixed-integer program.
2015	Hadi Vencheh <i>et al.</i> [34]	Suggests an inverse DEA model that replaces fluctuating DMUs with modified counterparts. Additionally, introduces an imprecise inverse DEA model designed for assessing relative efficiency when dealing with interval data.
2015	Azizi <i>et al.</i> [35]	Presents two approaches for obtaining efficiency scores in the presence of imprecise data using the SBM model: one optimistic (best-case) perspective and one pessimistic (worst-case) perspective.
2016	He <i>et al.</i> [36]	Develops interval efficiencies for inefficient DMUs through the utilization of interval ideal points, allowing for an increase in the lower bound of efficiency while enabling the upper bound of efficiency to reach unity (one).
2017	Shirdel <i>et al.</i> [37]	Proposes a model for assessing DMUs using interval data and applies its outcomes to rank the DMUs.

TABLE 1. continued.

Publication date	Author(s)	Main contribution to the literature
2017	Wei and Wang [38]	Creates a robust efficiency analysis procedure for cases with incomplete data, and introduces methods for perfect and potential efficiency analysis to establish lower and upper bounds for robust efficiency ratings.
2018	Toloo <i>et al.</i> [39]	Suggests a pair of interval DEA models for calculating interval efficiencies, considering inputs, outputs, and dual-role factors as intervals. These models utilize optimal multipliers for each dual-role factor to determine whether it functions as an input, output, or remains in equilibrium.
2020	Mo <i>et al.</i> [40]	Introduces an Interval Modified Slack-Based Measure (SBM) model designed for situations involving undesirable outputs.
2021	Ghobadi [41]	Suggests solutions to address challenges in merging DMUs when dealing with interval data.

A significant amount of literature has been dedicated to the study of two-stage DEA models, which can be considered a specialized type of the series mode. Sexton and Lewis [61] developed a model and used it to evaluate the efficiency of Major League Baseball. Kao and Hwang [62] proposed a method to decompose the efficiency of two-stage structures. Chen *et al.* [63] developed a similar efficiency decomposition method for variable returns to scale (VRS). Liang *et al.* [64] introduced efficiency measures for two-stage structures, including feedback loops where the output of the second stage may be used as input for the first stage. Liu and Lu [65] used the concept of eigenvector centrality in social network analysis to identify the efficient DMUs in two-stage systems. Halkos *et al.* [66] used an additive efficiency decomposition method and proposed a weight assurance region model to restrict the weights of the relative importance of the first and second stage performance. Aviles-Sacoto *et al.* [67] considered a situation where some outputs from the first stage may also be considered final outputs and developed a methodology to handle this situation. Tavana *et al.* [68] presented a two-stage DEA method for evaluating the performance of multi-level supply chains. Azizi and Kazemi-Matin [69] proposed a method for ranking two-stage production systems based on the efficiency score of each stage. Halkos *et al.* [70] provided a classification of the literature on two-stage DEA structures. Finally, Shahbazifar *et al.* [71] introduced a new model for evaluating the group efficiency of two-stage production systems.

In NDEA, the mixed or general state combines the series and parallel modes, where a process can use input from other processes and produce intermediate or final outputs. Limited research has been conducted on modeling the general structure of NDEA. Färe and Grosskopf [42] coined the term NDEA and developed an envelopment model for certain forms of network mixed states. Lewis and Sexton [72] extended the definition of NDEA and applied it to evaluate the American Baseball League. Kao [57] proposed a multiplier model for a mixed structure, and Kao and Hwang [73] explored further forms of mixed state and their impact on performance in the banking industry. Lozano [74] studied general network structure by defining a production possibility set (PPS) for each process and the entire system, and introduced a radial model for technical efficiency and another for cost efficiency. Kazemi Matin and Azizi [75] presented a model for network general state in both multiplier and envelopment forms, where a section of the process output could be used as input to other processes, and the remaining part of the output is the final output. Boloori *et al.* [76] proposed dual multiplier and envelopment DEA models for a more general state where a factor could be both an intermediate product and the main input-output simultaneously. Kalhor and Kazemi-Matin [77] introduced a new PPS for general network DEA in the presence of undesirable outputs and proposed network DEA models for performance evaluation of DMUs, emphasizing the use of distinct abatement factors and selecting an appropriate balance constraint. Yang *et al.* [78] proposed a two-level maximin NDEA model, considering the efficiency of all divisions and DMUs simultaneously.

The investigation of interval data in NDEA has been relatively limited. One of the few studies in this area was conducted by Amirteimoori and Kordrostami [14], who proposed a model for measuring the efficiency of a multi-component structure in a network parallel state with imprecise data. Kao and Liu [30] developed an approach for calculating interval efficiency in a two-stage network system with fuzzy data. They used a multiplier model to determine the upper bound of efficiency and a development model to calculate the lower bound of efficiency. Building on the work of Kao and Liu [30], Zhu and Zhou [32] explored the presence of imprecise data in a two-stage network structure.

While NDEA has seen limited exploration in the context of interval data, it is essential to examine its applications in various domains, including agriculture. The previous sections have laid the foundation by discussing the significance of DEA models in performance evaluation, the emergence of interval data analysis as a promising technique, and the categorization of NDEA models based on the internal structure of DMUs. In this section, we delve deeper into the intersection of interval data and NDEA, aiming to bridge the gap between these domains and extend the practical utility of NDEA models. Our investigation not only contributes to the theoretical understanding of interval data in network DEA but also sets the stage for its real-world application, as demonstrated in subsequent sections.

The field of agricultural economics has a profound influence on food security and economic stability, making it a prime arena for the application of DEA models. As modern agriculture becomes increasingly intricate, the necessity for advanced DEA models becomes apparent. These models are instrumental in addressing complex challenges such as scale inefficiency, environmental sustainability, and technological progress. Recent systematic reviews in the agricultural DEA domain have shed light on the shortcomings of traditional approaches like Constant Returns to Scale (CRS) and VRS, underscoring the demand for more sophisticated techniques. Remarkably, despite its paramount significance, NDEA models remain relatively underexplored in the agricultural sector, as noted by Kyrgiakos *et al.* [79].

Furthermore, Network DEA has been deployed in several studies to uncover inefficiencies within various subsystems of the agricultural process. Saputri *et al.* [80] employed this methodology to evaluate the efficiency of three distinct stages in the agri-food supply chain for Indonesian rice producers. Kord *et al.* [81] adopted a similar approach, presenting agricultural activity as two separate stages (environmental and economic) and conducting a sustainability assessment for different Iranian regions by leveraging shared inputs between these stages. Lu *et al.* [82] introduced a three-stage Network model to assess agricultural food production systems in European Union countries under the principles of a circular economy, where the final output carries over into the next period.

Additionally, Kord *et al.* [83] incorporated sensitivity analysis into their approach to evaluate the allocation of human resources in a two-stage Network DEA model. Specifically, they examined the optimal allocation of human resources in the initial cultivation and maintenance stage, as well as in the harvesting stage. Abbas *et al.* [84] also utilized this analysis to investigate changes in crop output under varying input conditions. Furthermore, Grey Relational Analysis was applied by Yang *et al.* [85] to assess the impact of included variables on the environmental performance of Chinese households.

In light of these developments, it is worth noting that there is a noticeable gap in research that combines NDEA and agriculture, especially when dealing with interval data. Recognizing this research vacuum, our study aims to fill this void by presenting a practical application that focuses on evaluating wheat farming practices across 30 Iranian provinces. In doing so, our objective is to showcase the potential of NDEA models in enhancing the efficiency and sustainability of agricultural operations, thus pushing the boundaries of knowledge in the assessment of agricultural efficiency.

This study addresses the research gap in the limited number of papers that examine the presence of interval data in network DEA. The contribution of this study is twofold: firstly, it provides a new tool for evaluating network production processes with imprecise data, and secondly, it demonstrates the practical application of this tool in the agriculture sector.

The subsequent section of this paper examines the impact of interval data on general network structures of DEA and proposes novel models for assessing the upper and lower bounds of efficiency. In Section 3, an

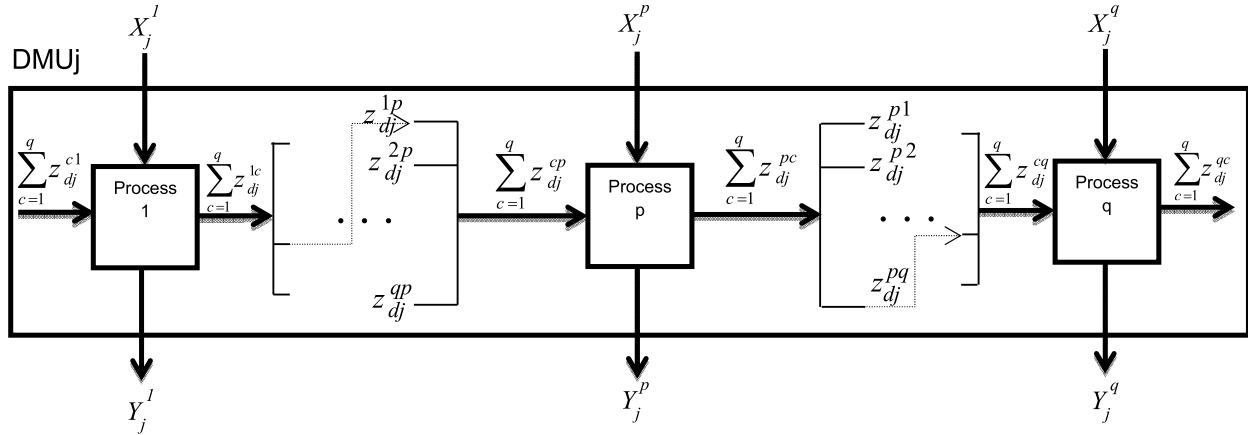


FIGURE 1. A production series network in a general operational mode.

illustrative example is formulated to demonstrate the effectiveness of the proposed models. Additionally, Section 4 showcases a real-world application of the models in evaluating wheat production in Iran. Finally, the study concludes with remarks presented in Section 5. In the following sections, we delve into the impact of interval data on general network structures, provide novel models for assessing efficiency bounds, present an illustrative example, discuss a real-world application in wheat production, and conclude with our findings and remarks.

2. INTERVAL DATA IN NDEA

As previously stated, NDEA can be classified into three main states based on the internal structure of DMUs, namely parallel, series, and mixed. Since the overall network structure encompasses all three states, this section focuses solely on the general states in relation to interval data.

2.1. General network structure with interval data

In a general state, each process has the ability to consume exogenous inputs and intermediate products produced by other processes as inputs, and produce final outputs and other intermediate products as outputs. Let us consider n observed DMUs. Figure 1 illustrates the internal relationships among processes for the production units in the general mode. Specifically, the figure represents DMU_j, where x_{ij} ; $i = 1, \dots, m$, refers to the observed amount of the i -th exogenous input to that DMU. Each DMU is divided into q processes, labelled as p ; $p = 1, \dots, q$. The amount of the i -th input consumed by the p -th process in DMU_j is indicated by x_{ij}^p , representing the total amount of the i -th input consumed by all processes in DMU_j ($x_{ij} = \sum_{p=1}^q x_{ij}^p$). Similarly, y_{rj} ; $r = 1, \dots, s$, denotes the amount of the r -th final output of the DMU, while y_{rj}^p represents the amount of the r -th final output produced by the p -th process in DMU_j. The flow of outputs among the processes is represented by $y_{rj} = \sum_{p=1}^q y_{rj}^p$. Additionally, intermediate products are generated and consumed by the various processes within the DMU.

Let z_{dj}^{pc} denote the amount of the d -th intermediate product that is produced by the p -th process and consumed by the c -th process in DMU_j. We define $\sum_{c=1}^q z_{dj}^{pc}$ as the total amount of the d -th intermediate product produced by the p -th process in DMU_j, and $\sum_{c=1}^q z_{dj}^{cp}$ as the total amount of the d -th intermediate product consumed by the p -th process in DMU_j.

The Model (1) proposed by Kazemi-Matin and Azizi [75] is used to evaluate the efficiency score of DMU_k for relations and internal structures similar to those depicted in Figure 1:

$$\begin{aligned}
 \theta_k = \text{Max} \quad & \sum_{p=1}^q \sum_{r=1}^s u_r y_{rk}^p \\
 \text{s.t.} \quad & \sum_{p=1}^q \sum_{i=1}^m v_i x_{ik}^p = 1 \\
 & \sum_{r=1}^s u_r y_{rj}^p - \sum_{i=1}^m v_i x_{ij}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p z_{dj}^{pc} - \sum_{c=1}^q \sum_{d=1}^D w_d^c z_{dj}^{cp} \leq 0, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & v_i \geq \varepsilon \quad i = 1, \dots, m \\
 & u_r \geq \varepsilon \quad r = 1, \dots, s \\
 & w_d^p \geq \varepsilon \quad d = 1, \dots, D, \quad p = 1, \dots, q. \quad (1)
 \end{aligned}$$

If we define the inputs, outputs, and intermediate products as intervals $\hat{x}_{ij}^p \in [\underline{x}_{ij}^p, \bar{x}_{ij}^p]$, $\hat{y}_{rj}^p \in [\underline{y}_{rj}^p, \bar{y}_{rj}^p]$ and $\hat{z}_{dj}^{pc} \in [\underline{z}_{dj}^{pc}, \bar{z}_{dj}^{pc}]$, respectively, then Model (1) becomes nonlinear. To obtain a linear version of the model, variable alternation and transformation can be applied as recommended by Despotis and Smirlis [5]. Specifically, a convex combination of the lower and upper bounds of each input, output, and intermediate product interval is used. To illustrate this approach, let's consider these equations:

$$\begin{aligned}
 \hat{x}_{ij}^p &= \underline{x}_{ij}^p + \alpha_{ij}^p (\bar{x}_{ij}^p - \underline{x}_{ij}^p); \quad 0 \leq \alpha_{ij}^p \leq 1 \\
 \hat{y}_{rj}^p &= \underline{y}_{rj}^p + \beta_{rj}^p (\bar{y}_{rj}^p - \underline{y}_{rj}^p); \quad 0 \leq \beta_{rj}^p \leq 1 \\
 \hat{z}_{dj}^{pc} &= \underline{z}_{dj}^{pc} + \gamma_{dj}^{pc} (\bar{z}_{dj}^{pc} - \underline{z}_{dj}^{pc}); \quad 0 \leq \gamma_{dj}^{pc} \leq 1.
 \end{aligned}$$

To transform Model (1) into a linear model (Model (2)), we make variable substitutions: $q_{ij}^p = v_i \alpha_{ij}^p$, $p_{rj}^p = u_r \beta_{rj}^p$, and $t_{dj}^{pc} = w_d^p \gamma_{dj}^{pc}$, and introduce new decision variables q_{ij}^p , p_{rj}^p and t_{dj}^{pc} . This results in the following linear model for evaluating the efficiency scores of DMU_k :

$$\begin{aligned}
 \theta_k = \text{Max} \quad & \sum_{p=1}^q \sum_{r=1}^s u_r y_{rk}^p + p_{rk}^p (\bar{y}_{rk}^p - \underline{y}_{rk}^p) \\
 \text{s.t.} \quad & \sum_{p=1}^q \sum_{i=1}^m v_i \underline{x}_{ik}^p + q_{ik}^p (\bar{x}_{ik}^p - \underline{x}_{ik}^p) = 1 \\
 & \sum_{r=1}^s u_r \underline{y}_{rj}^p + p_{rj}^p (\bar{y}_{rj}^p - \underline{y}_{rj}^p) - \sum_{i=1}^m v_i \underline{x}_{ij}^p + q_{ij}^p (\bar{x}_{ij}^p - \underline{x}_{ij}^p) \\
 & + \sum_{c=1}^q \sum_{d=1}^D w_d^p \underline{z}_{dj}^{pc} + t_{dj}^{pc} (\bar{z}_{dj}^{pc} - \underline{z}_{dj}^{pc}) - \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dj}^{cp} \\
 & + t_{dj}^{cp} (\bar{z}_{dj}^{cp} - \underline{z}_{dj}^{cp}) \leq 0, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & 0 \leq p_{rj}^p \leq u_r, \quad r = 1, \dots, s, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & 0 \leq q_{ij}^p \leq v_i, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & 0 \leq t_{dj}^{pc} \leq w_d^p, \quad d = 1, \dots, D, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & v_i \geq \varepsilon, \quad i = 1, \dots, m
 \end{aligned}$$

$$\begin{aligned} u_r &\geq \varepsilon, & r &= 1, \dots, s \\ w_d^p &\geq \varepsilon, & d &= 1, \dots, D. \end{aligned} \quad (2)$$

The above variable substitutions are made to transform the original nonlinear model into a linear model that can be easily solved using linear programming techniques. In addition, the new decision variables are introduced to represent the deviation of the actual input/output values from the lower bound of the corresponding intervals. By using a convex combination of the lower and upper bounds of each input, output, and intermediate product interval, we ensure that the resulting linear model is equivalent form of the original nonlinear model.

Solving the linear model (Model (2)) yields efficiency scores for each DMU_k, which indicate how well the DMU_k utilizes its inputs to generate outputs relative to the other DMUs in the sample. These efficiency scores can be used to identify best practices and benchmarking targets for improving the efficiency of the DMUs under evaluation.

To reduce the number of variables and computation cost while obtaining upper and lower bounds of efficiency scores in DEA with interval data, the strategy is to use the principle of convexity. For a specific DMU, the inputs are assumed to be at the lower bound and the outputs at the upper bound within the intervals to determine the highest efficiency score. The worst-case scenario within the intervals is selected for both inputs and outputs of the other units in this case. Conversely, for the lowest efficiency score of the same DMU, the inputs are assumed to be at the upper bound and the outputs at the lower bound, while the best-case scenario within the intervals is selected for both inputs and outputs of the other units. This approach allows for the computation of the upper and lower bounds of the efficiency score without calculating the actual efficiency score for each DMU, which is challenging when dealing with large datasets. It should be noted that since intermediate products serve as both the outputs of the previous stage and the inputs of the next stage, the model can select the best values for them.

In Model (2), the DMU_k achieves its highest efficiency levels when its outputs reach their maximum values $\forall r : \hat{y}_{rk} = \bar{y}_{rk}$ and its inputs reach their minimum values $\forall i : \hat{x}_{ik} = \underline{x}_{ik}$, at the same time. However, the outputs of the remaining DMUs reach their minimum values while their inputs reach their maximum levels, *i.e.* $\forall r \forall j \neq k : \hat{y}_{rj} = \underline{y}_{rj}$ and $\forall i \forall j \neq k : \hat{x}_{ij} = \bar{x}_{ij}$; indicating that these DMUs achieve their highest values of inputs and lowest values of outputs. Conversely, DMU_k reaches its lowest efficiency level when its outputs reach their minimum values and inputs reach their maximum values, while the remaining DMUs achieve their highest efficiency levels with their outputs reaching their maximum values and inputs reaching their minimum values. This implies that the remaining DMUs achieve their lowest values of inputs and highest values of outputs.

$$\begin{aligned} \forall p \forall r : p_{rj}^p &= \begin{cases} u_r & j = k \\ 0 & j \neq k \end{cases} \\ \forall p \forall i : q_{ij}^p &= \begin{cases} 0 & j = k \\ v_i & j \neq k. \end{cases} \end{aligned}$$

Intermediate products generated by a process must be entirely utilized by other processes, meaning that altering the value of an intermediate product from one process has a ripple effect on the other processes. If the value of an intermediate product produced by process p rises, the input values of other processes will also rise, leading to a decrease in the efficiencies of those processes. Therefore, intermediate products cannot be treated as independent variables since changes in their values have an impact on the efficiencies of the other processes.

Similarly, the same principle can be applied to intermediate products that are consumed by process p . These intermediate products must have been influenced by other processes as well. Thus, if we reduce the value of an intermediate product produced by process p to improve its efficiency, it will also decrease the output values of other processes. Consequently, the efficiencies of these processes will decrease as well. Therefore, we cannot consider the lowest values of the consumed intermediate products as independent variables.

Model (3) is proposed below to calculate the highest efficiency level of DMU_k :

$$\begin{aligned}
 \bar{\theta}_k = \text{Max} \quad & \sum_{p=1}^q \sum_{r=1}^s u_r \bar{y}_{rk}^p \\
 \text{s.t.} \quad & \sum_{p=1}^q \sum_{i=1}^m v_i \underline{x}_{ik}^p = 1 \\
 & \sum_{r=1}^s u_r \bar{y}_{rk}^p - \sum_{i=1}^m v_i \underline{x}_{ik}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p z_{dk}^{pc} + t_{dk}^{pc} (\bar{z}_{dk}^{pc} - \underline{z}_{dk}^{pc}) \\
 & - \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dk}^{cp} + t_{dk}^{cp} (\bar{z}_{dk}^{cp} - \underline{z}_{dk}^{cp}) \leq 0, \quad p = 1, \dots, q \\
 & \sum_{r=1}^s u_r \bar{y}_{rj}^p - \sum_{i=1}^m v_i \bar{x}_{ij}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p z_{dj}^{pc} + t_{dj}^{pc} (\bar{z}_{dj}^{pc} - \underline{z}_{dj}^{pc}) \\
 & - \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dj}^{cp} + t_{dj}^{cp} (\bar{z}_{dj}^{cp} - \underline{z}_{dj}^{cp}) \leq 0, \quad p = 1, \dots, q, \quad j = 1, \dots, n; \quad j \neq k \\
 & 0 \leq t_{dj}^{pc} \leq w_d^p, \quad d = 1, \dots, D, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & v_i \geq \varepsilon, \quad i = 1, \dots, m \\
 & u_r \geq \varepsilon, \quad r = 1, \dots, s \\
 & w_d^p \geq \varepsilon, \quad d = 1, \dots, D. \quad (3)
 \end{aligned}$$

Theorem 1. Optimal values of models (2) and (3) are equal.

Proof. This theorem has been proven in the ratio model of evaluating efficiency scores. It has been shown that Models (2) and (3) are equivalent to Models (2') and (3'), respectively, as follows:

$$\begin{aligned}
 \theta_k = \text{Max} \quad & \frac{\sum_{p=1}^q \sum_{r=1}^s u_r \bar{y}_{rk}^p + p_{rk}^p (\bar{y}_{rk}^p - \underline{y}_{rk}^p)}{\sum_{p=1}^q \sum_{i=1}^m v_i \underline{x}_{ik}^p + q_{ik}^p (\bar{x}_{ik}^p - \underline{x}_{ik}^p)} \\
 & \frac{\sum_{r=1}^s u_r \bar{y}_{rj}^p + p_{rj}^p (\bar{y}_{rj}^p - \underline{y}_{rj}^p)}{\sum_{i=1}^m v_i \underline{x}_{ij}^p + q_{ij}^p (\bar{x}_{ij}^p - \underline{x}_{ij}^p)} \\
 & + \frac{\sum_{c=1}^q \sum_{d=1}^D w_d^p z_{dj}^{pc} + t_{dj}^{pc} (\bar{z}_{dj}^{pc} - \underline{z}_{dj}^{pc})}{\sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dj}^{cp} + t_{dj}^{cp} (\bar{z}_{dj}^{cp} - \underline{z}_{dj}^{cp})} \leq 1, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & 0 \leq p_{rj}^p \leq u_r, \quad r = 1, \dots, s, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & 0 \leq q_{ij}^p \leq v_i, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & 0 \leq t_{dj}^{pc} \leq w_d^p, \quad d = 1, \dots, D, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
 & v_i \geq \varepsilon, \quad i = 1, \dots, m \\
 & u_r \geq \varepsilon, \quad r = 1, \dots, s \\
 & w_d^p \geq \varepsilon, \quad d = 1, \dots, D \quad (2')
 \end{aligned}$$

$$\begin{aligned}
\bar{\theta}_k &= \text{Max} \frac{\sum_{p=1}^q \sum_{r=1}^s u_r \bar{y}_{rk}^p}{\sum_{p=1}^q \sum_{i=1}^m v_i \underline{x}_{ik}^p} \\
\text{s.t. } &\frac{\sum_{r=1}^s u_r \bar{y}_{rk}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p \bar{z}_{dk}^{pc} + t_{dk}^{pc} (\bar{z}_{dk}^{pc} - \underline{z}_{dk}^{pc})}{\sum_{i=1}^m v_i \underline{x}_{ik}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dk}^{cp} + t_{dk}^{cp} (\bar{z}_{dk}^{cp} - \underline{z}_{dk}^{cp})} \leq 1, \quad p = 1, \dots, q \\
&\frac{\sum_{r=1}^s u_r \bar{y}_{rj}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p \bar{z}_{dj}^{pc} + t_{dj}^{pc} (\bar{z}_{dj}^{pc} - \underline{z}_{dj}^{pc})}{\sum_{i=1}^m v_i \bar{x}_{ij}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dj}^{cp} + t_{dj}^{cp} (\bar{z}_{dj}^{cp} - \underline{z}_{dj}^{cp})} \leq 1, \quad p = 1, \dots, q, \quad j = 1, \dots, n, \quad j \neq k \\
&0 \leq t_{dj}^{pc} \leq w_d^p, \quad d = 1, \dots, D, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
&v_i \geq \varepsilon, \quad i = 1, \dots, m \\
&u_r \geq \varepsilon, \quad r = 1, \dots, s \\
&w_d^p \geq \varepsilon, \quad d = 1, \dots, D. \tag{3'}
\end{aligned}$$

Let $(U^*, V^*, W^*, P^*, Q^*, T^*)$ be the optimal solution for Model (2') with optimal value of θ_k^* , where $U^* = (u_r^*; r = 1, \dots, s)$, $V^* = (v_i^*; i = 1, \dots, m)$, $W^* = (w_d^p; d = 1, \dots, D, p = 1, \dots, q)$, $P^* = (p_{rj}^*; r = 1, \dots, s, j = 1, \dots, n, p = 1, \dots, q)$, $Q^* = (q_{ij}^*; i = 1, \dots, m, j = 1, \dots, n, p = 1, \dots, q)$ and $T^* = (t_{dj}^{pc}; d = 1, \dots, D, j = 1, \dots, n, p, c = 1, \dots, q)$. Therefore,

$$\begin{aligned}
1 &\geq \frac{\sum_{r=1}^s u_r^* \bar{y}_{rj}^p + p_{rj}^* (\bar{y}_{rj}^p - \underline{y}_{rj}^p) + \sum_{c=1}^q \sum_{d=1}^D w_d^p \bar{z}_{dj}^{pc} + t_{dj}^{pc} (\bar{z}_{dj}^{pc} - \underline{z}_{dj}^{pc})}{\sum_{i=1}^m v_i^* \underline{x}_{ij}^p + q_{ij}^* (\bar{x}_{ij}^p - \underline{x}_{ij}^p) + \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dj}^{cp} + t_{dj}^{cp} (\bar{z}_{dj}^{cp} - \underline{z}_{dj}^{cp})} \\
&\geq \frac{\sum_{r=1}^s u_r^* \bar{y}_{rj}^p + 0(\bar{y}_{rj}^p - \underline{y}_{rj}^p) + \sum_{c=1}^q \sum_{d=1}^D w_d^p \bar{z}_{dj}^{pc} + t_{dj}^{pc} (\bar{z}_{dj}^{pc} - \underline{z}_{dj}^{pc})}{\sum_{i=1}^m v_i^* \underline{x}_{ij}^p + v_i^* (\bar{x}_{ij}^p - \underline{x}_{ij}^p) + \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dj}^{cp} + t_{dj}^{cp} (\bar{z}_{dj}^{cp} - \underline{z}_{dj}^{cp})} \\
&= \frac{\sum_{r=1}^s u_r^* \bar{y}_{rj}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p \bar{z}_{dj}^{pc} + t_{dj}^{pc} (\bar{z}_{dj}^{pc} - \underline{z}_{dj}^{pc})}{\sum_{i=1}^m v_i^* \bar{x}_{ij}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dj}^{cp} + t_{dj}^{cp} (\bar{z}_{dj}^{cp} - \underline{z}_{dj}^{cp})}, \quad j = 1, \dots, n, \quad p = 1, \dots, q.
\end{aligned}$$

The optimal value of Model (2)'s objective function increases consistently as Q decreases and as P increases. This means that the optimal value of P should be the highest possible, denoted as $P^* = U^*$, while the optimal value of Q should be the lowest possible, represented as $Q^* = 0$. As a result, the k th constraints of Model (2) can be modified accordingly:

$$\begin{aligned}
&\frac{\sum_{r=1}^s u_r^* \bar{y}_{rk}^p + u_r^* (\bar{y}_{rk}^p - \underline{y}_{rk}^p) + \sum_{c=1}^q \sum_{d=1}^D w_d^p \bar{z}_{dk}^{pc} + t_{dk}^{pc} (\bar{z}_{dk}^{pc} - \underline{z}_{dk}^{pc})}{\sum_{i=1}^m v_i^* \underline{x}_{ik}^p + 0(\bar{x}_{ik}^p - \underline{x}_{ik}^p) + \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dk}^{cp} + t_{dk}^{cp} (\bar{z}_{dk}^{cp} - \underline{z}_{dk}^{cp})} \leq 1 \quad p = 1, \dots, q \\
&\Rightarrow \frac{\sum_{r=1}^s u_r^* \bar{y}_{rk}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p \bar{z}_{dk}^{pc} + t_{dk}^{pc} (\bar{z}_{dk}^{pc} - \underline{z}_{dk}^{pc})}{\sum_{i=1}^m v_i^* \underline{x}_{ik}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^c \underline{z}_{dk}^{cp} + t_{dk}^{cp} (\bar{z}_{dk}^{cp} - \underline{z}_{dk}^{cp})} \leq 1 \quad p = 1, \dots, q.
\end{aligned}$$

Because (U^*, V^*, W^*, T^*) is a feasible solution for Model (3'), then $\theta_k^* \leq \bar{\theta}_k^*$. Conversely, let (U^*, V^*, W^*, T^*) be the optimal solution of Model (3') with optimal value $\bar{\theta}_k^*$. Defining $\bar{P}^* = (\bar{p}_{rj}^*; r = 1, \dots, s, j = 1, \dots, n, p = 1, \dots, q)$ and $\bar{Q}^* = (\bar{q}_{ij}^*; i = 1, \dots, m, j = 1, \dots, n, p = 1, \dots, q)$ so that $\bar{p}_{rj}^p = \begin{cases} u_r^*; j = k \\ 0; j \neq k \end{cases}$ and $\bar{q}_{ij}^p = \begin{cases} 0; j = k \\ v_i^*; j \neq k \end{cases}$, then $(U^*, V^*, W^*, P^*, Q^*, T^*)$ is a feasible solution for Model (2') concluding that $\theta_k^* \geq \bar{\theta}_k^*$. This completes the proof. $\square$

Therefore, to achieve the highest total efficiency, the optimal values of intermediate products should be treated as variables within the interval of $[\underline{z}_{dj}^{pc}, \bar{z}_{dj}^{pc}]$. Thus, a convex combination of intervals should be utilized when modelling to obtain the upper bound of efficiency of intermediate products.

Note that Model (2) is designed to incorporate all data levels within the intervals, whereas Model (3) acts as its linear counterpart tailored for precise data. Notably, we demonstrate that Model (2) can be seamlessly converted into Model (3), enabling the computation of upper bounds on efficiency scores in scenarios with imprecise data. This transformation significantly bolsters practicality and computational efficiency, rendering Model (3) particularly well-suited for large-scale applications and real-time decision-making.

Consequently, DMU_k achieves the lowest efficiency level in the most unfavorable state, where the outputs of each process of DMU_k reach their minimum values while the inputs of each process of DMU_k reach their maximum values simultaneously. Meanwhile, the remaining DMUs' processes reach their maximum output values and minimum input values simultaneously. As we discussed earlier, to determine the optimal values of intermediate products that lead to the lowest efficiency, they should be considered as variables within their intervals. Thus, an optimal solution for the following Model (4) would indicate the lowest efficiency level for DMU_k :

$$\begin{aligned}
\theta_k = & \text{Max} \sum_{p=1}^q \sum_{r=1}^s u_r y_{rk}^p \\
\text{s.t.} \quad & \sum_{p=1}^q \sum_{i=1}^m v_i \bar{x}_{ik}^p = 1 \\
& \sum_{r=1}^s u_r \bar{y}_{rk}^p - \sum_{i=1}^m v_i \bar{x}_{ik}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p z_{dk}^{pc} + t_{dk}^{pc} (\bar{z}_{dk}^{pc} - z_{dk}^{pc}) \\
& - \sum_{c=1}^q \sum_{d=1}^D w_d^c z_{dk}^{cp} + t_{dk}^{cp} (\bar{z}_{dk}^{cp} - z_{dk}^{cp}) \leq 0, \quad p = 1, \dots, q \\
& \sum_{r=1}^s u_r \bar{y}_{rj}^p - \sum_{i=1}^m v_i \bar{x}_{ij}^p + \sum_{c=1}^q \sum_{d=1}^D w_d^p z_{dj}^{pc} + t_{dj}^{pc} (\bar{z}_{dj}^{pc} - z_{dj}^{pc}) \\
& - \sum_{c=1}^q \sum_{d=1}^D w_d^c z_{dj}^{cp} + t_{dj}^{cp} (\bar{z}_{dj}^{cp} - z_{dj}^{cp}) \leq 0, \quad p = 1, \dots, q, \quad j = 1, \dots, n; \quad j \neq k \\
& 0 \leq t_{dj}^{pc} \leq w_d^p, \quad d = 1, \dots, D, \quad j = 1, \dots, n, \quad p = 1, \dots, q \\
& v_i \geq \varepsilon, \quad i = 1, \dots, m \\
& u_r \geq \varepsilon, \quad r = 1, \dots, s \\
& w_d^p \geq \varepsilon, \quad d = 1, \dots, D.
\end{aligned} \tag{4}$$

The optimal value obtained through Model (4) can be regarded as the lowest boundary for efficiency scores of DMU_k .

2.2. Classifying DMUs

After solving the models for the upper and lower boundaries of efficiency, DMUs can be categorized into one of the following four groups:

$$\begin{aligned}
E^{++} &= \left\{ DMU_k | \theta_k \geq \max_{j \neq k} \bar{\theta}_j \right\} \\
E^+ &= \left\{ DMU_k | \theta_k = \max_j \underline{\theta}_j \quad \text{or} \quad \bar{\theta}_k = \max_j \bar{\theta}_j \right\} \\
E^- &= \left\{ DMU_k | \theta_k < \max_{j \neq k} \underline{\theta}_j \quad \text{and} \quad \bar{\theta}_k < \max_{j \neq k} \bar{\theta}_j \right\} \\
E^{--} &= \left\{ DMU_k | \bar{\theta}_k \leq \min_{j \neq k} \underline{\theta}_j \right\}.
\end{aligned} \tag{5}$$

After solving the DEA models for the upper and lower boundaries of efficiency, DMUs can be classified into one of four groups: E^{++} , E^+ , E^- , and E^{--} . These groups are defined based on the upper and lower bounds of efficiency scores of the DMUs relative to the other DMUs in the network.

The first group, E^{++} , consists of DMUs whose worst efficiency score is greater than or equal to the best observed efficiency score of all other DMUs in the network. These DMUs are fully efficient in both the upper and lower bounds and represent the upper boundary of the production frontier. They are the most efficient units in the network.

The second group, E^+ , includes DMUs that have an efficiency score equal to the maximum observed efficiency score on at least one of the efficiency bounds when compared to other DMUs in the network. These DMUs are efficient, but they may not be the most efficient in the network. This group encompasses DMUs that are on the frontier as well as those that are not but are still relatively efficient.

The third group, E^- , comprises inefficient DMUs that are not the best in either the upper or lower bounds. Specifically, a DMU is in this group if its efficiency score in both the upper and lower bounds is less than the efficiency scores of all other DMUs.

The fourth group, E^{--} , includes DMUs that are fully inefficient in both upper and lower bounds and are dominated by other DMUs in the dataset. In other words, their efficiency scores in both bounds are less than or equal to the efficiency scores of all other DMUs.

In summary, categorizing DMUs into E^{++} , E^+ , E^- , and E^{--} groups provides a way to identify the most efficient, efficient, inefficient, and fully inefficient units in a network. This information can be used to develop strategies for improving the efficiency of DMUs in the network. For example, DMUs in the E^{++} group may be considered as benchmarks for other DMUs to strive towards, while DMUs in the E^+ group may require further investigation to determine why they are only efficient in one bound. On the other hand, DMUs in the E^- and E^{--} groups may need improvement to increase their efficiency scores.

2.3. Efficiency improvement and target setting

To improve the efficiency of DMUs in the E^- or E^+ groups, we need to identify the sources of inefficiency by examining the lower boundary of their efficiency scores. The lower boundary indicates the minimum efficiency level that the DMUs can achieve, and the difference between the lower boundary and the highest efficiency score shows the amount of inefficiency. For these units we need to optimize the use of inputs by identifying ways to reduce input usage without affecting the quantity of the produced outputs. This can be achieved through the following target settings.

Let's DMU_k : $(\mathbf{x}_k, \mathbf{z}_k, \mathbf{y}_k)$ belongs to E^{--} . So, in evaluating this unit with Model (3) we get $\bar{\theta}_k < \min_{j \neq k} \underline{\theta}_j < 1$, then the new activity $(\bar{\theta}_k \mathbf{x}_k, \mathbf{z}_k, \mathbf{y}_k)$ is considered as a benchmark for this unit. A similar improvement scenario can also be implemented for units in the E^- and E^+ group.

Theorem 2. If $\text{DMU}_k(\mathbf{x}_k, \mathbf{z}_k, \mathbf{y}_k) \in E^-$, then $(\bar{\theta}_k \mathbf{x}_k, \mathbf{z}_k, \mathbf{y}_k) \in E^+ \cup E^{++}$.

Proof. It is enough to show that for the improved activity is efficient in Model (3). If we assume that the data are exact, by taking dual for the linear Model (1) and by removing the slack variables from the objective function and the constraints we have the following envelopment model which is equivalent to the multiplier Model (1) when the variables are non-negative:

$$\begin{aligned} & \text{Min } h_k \\ & \text{s.t. } \sum_{p=1}^q \sum_{j=1}^n \lambda_j^p x_{ij}^p \leq h \sum_{p=1}^q x_{ik}^p; \quad i = 1, \dots, m \\ & \quad \sum_{p=1}^q \sum_{j=1}^n \lambda_j^p y_{rj}^p \geq \sum_{p=1}^q y_{rk}^p; \quad r = 1, \dots, s \end{aligned}$$

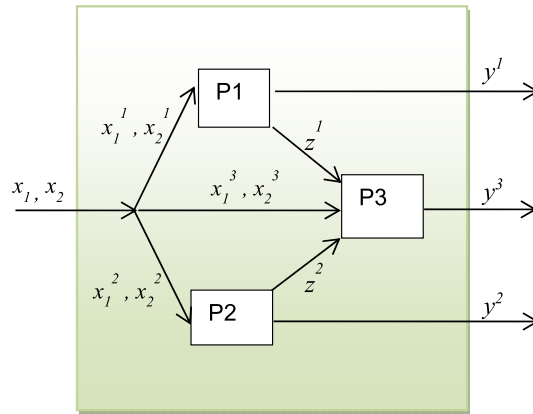


FIGURE 2. Internal structure of a DMU in illustrative example.

$$\begin{aligned} \sum_{c=1}^q \sum_{j=1}^n \lambda_j^p z_{dj}^{pc} &\geq \sum_{c=1}^q \sum_{j=1}^n \lambda_j^c z_{dj}^{cp}; & d = 1, \dots, D, \quad p = 1, \dots, q \\ \lambda_j^p &\geq 0; & j = 1, \dots, n, \quad p = 1, \dots, q. \end{aligned} \quad (D1)$$

It is easy to see that this model is always feasible and in optimality $0 < h_k^* \leq 1$. Note that for calculating upper bound of the efficiency for DMU_k, this unit is set at its best level while the other units are at their worst level and the computed score is $\bar{\theta}_k < \max_j \theta_j \leq 1$. Now consider the interval data when the new activity $(\bar{\theta}_k \hat{\mathbf{x}}_k, \hat{\mathbf{z}}_k, \hat{\mathbf{y}}_k)$ is under evaluation for its upper bound of the efficiency score. We show that the new upper bound computed by the following model is equal to $\max_j \bar{\theta}_j = 1$:

$$\begin{aligned} &\text{Min } h_k \\ &\text{s.t. } \sum_{p=1}^q \sum_{j=1}^n \lambda_j^p \bar{x}_{ij}^p \leq h_k \sum_{p=1}^q \bar{\theta}_k \underline{x}_{ik}^p; & i = 1, \dots, m \\ &\quad \sum_{p=1}^q \sum_{j=1}^n \lambda_j^p \underline{y}_{rj}^p \geq \sum_{p=1}^q \bar{y}_{rk}^p; & r = 1, \dots, s \\ &\quad \sum_{c=1}^q \sum_{j=1}^n \lambda_j^p z_{dj}^{pc} \geq \sum_{c=1}^q \sum_{j=1}^n \lambda_j^c z_{dj}^{cp}; & d = 1, \dots, D, \quad p = 1, \dots, q \\ &\quad \lambda_j^p \geq 0; & j = 1, \dots, n, \quad p = 1, \dots, q. \end{aligned}$$

For simplicity of presentation the intermediate variables $\hat{z}_{dj}^{pc}$ and $\hat{z}_{dj}^{cp}$ are stated as interval values. If we consider the first constraint as $\sum_{p=1}^q \sum_{j=1}^n \lambda_j^p \bar{x}_{ij}^p \leq \sum_{p=1}^q (h_k \bar{\theta}_k) \underline{x}_{ik}^p$, it is easy to see that the optimal solution $(h_k^* \bar{\theta}_k, \lambda_j^{*p} \geq 0; j = 1, \dots, n, p = 1, \dots, q)$ is a feasible solution for the model when calculating upper bound of the efficiency for DMU_k. Since $\bar{\theta}_k$ is the optimal upper value we have $\bar{\theta}_k \leq h_k^* \bar{\theta}_k$ which leads to $1 \leq h_k^*$ and we get $h_k^* = 1 = \max_j \bar{\theta}_j$. $\square$

3. ILLUSTRATIVE EXAMPLE

The present example's structure and data have been derived from Kao's [57] work. However, certain modifications have been made and the data has been assumed to be interval. It is important to note that the Kao [57] example's data has been utilized as the lower boundary for the interval data in the current instance. The current study examines five DMUs that exhibit a network structure comparable to that shown in Figure 2.

Each DMU consumes two inputs, namely x_1 and x_2 , which are common to all three processes. Three outputs, namely y_1 (generated by process 1), y_2 (generated by process 2), and y_3 (generated by process 3), are produced. Additionally, intermediate product z_1 is generated by process 1, intermediate product z_2 is generated by process 2, and both intermediate products are consumed by process 3. The data in this example is interval in nature and is presented in Table 2.

Due to the mixed (general) internal structure of the DMUs, it was necessary to modify Model (3) to obtain the upper boundary of the efficiency score in Model (6). The modification was carried out in the following manner:

$$\begin{aligned}
 \bar{\theta}_k &= \text{Max} \sum_{p=1}^3 u_p \bar{y}_k^p \\
 \text{s.t.} \quad & \sum_{p=1}^3 \sum_{i=1}^2 v_i \underline{x}_{ik}^p = 1 \\
 & u_p \bar{y}_k^p - \sum_{i=1}^2 v_i \underline{x}_{ik}^p + w^p \underline{z}_k^p + t_k^p (\bar{z}_k^p - \underline{z}_k^p) \leq 0, \quad p = 1, 2 \\
 & u_p \underline{y}_j^p - \sum_{i=1}^2 v_i \bar{x}_{ij}^p + w^p \underline{z}_j^p + t_j^p (\bar{z}_j^p - \underline{z}_j^p) \leq 0, \quad p = 1, 2, \quad j = 1, \dots, 5; \quad j \neq k \\
 & u_3 \bar{y}_k^3 - \sum_{i=1}^2 v_i \underline{x}_{ik}^3 - \sum_{p=1}^2 w^p \underline{z}_k^p + t_k^p (\bar{z}_k^p - \underline{z}_k^p) \leq 0 \\
 & u_3 \underline{y}_j^3 - \sum_{i=1}^2 v_i \bar{x}_{ij}^3 - \sum_{p=1}^2 w^p \underline{z}_j^p + t_j^p (\bar{z}_j^p - \underline{z}_j^p) \leq 0, \quad j = 1, \dots, 5; \quad j \neq k \\
 & 0 \leq t_j^p \leq w^p, \quad j = 1, \dots, 5, \quad p = 1, 2 \\
 & v_i \geq \varepsilon, \quad i = 1, 2 \\
 & u_p \geq \varepsilon, \quad p = 1, 2, 3 \\
 & w^p \geq \varepsilon, \quad p = 1, 2.
 \end{aligned} \tag{6}$$

The results obtained by applying Model (6) to the current scenario, while considering a value of $\varepsilon = 10^{-5}$, are presented in Table 3.

The determination of variables related to the convex combination of intermediate products can be achieved by acquiring knowledge of both w^p and t_j^p , where p denotes a particular product. The optimal values of intermediate products can be computed by utilizing γ_j^p in equation $\hat{z}_j^p = \gamma_j^c \bar{z}_j^p + (1 - \gamma_j^p) \underline{z}_j^p$. The results of this computation are presented in two columns of Table 3.

An interesting observation is that the highest efficiency score for DMUs does not necessarily correspond to the lowest or highest values of the intermediate product interval. Additionally, it is apparent that for DMU3, the highest efficiency score is achieved when the intermediate products of both DMU3 and DMU5 are situated within the range of their lowest and highest values. Likewise, the lower limit of efficiency scores can be determined for all five DMUs by utilizing Model (4), and the resulting values are displayed in the last column of Table 3.

Following the computation of upper and lower boundaries of the efficiency score, an interval efficiency score has been calculated for each DMU. As a result, DMUs 1, 2, and 5 have been identified as members of the set E^+ due to their complete efficiency scores being at the upper boundary, while DMUs 3 and 4 have been identified as members of the set E^- due to their efficiency scores being less than best observed boundary values. Both E^{++} and E^{--} groups are empty.

TABLE 2. Inputs, intermediate products and outputs for 5 DMUs with mixed structure relations in sub-processes.

DMUs	Processes	x_1	x_2	z^1	z^2	y^1	y^2	y^3
1	1	3	5	[2, 4]	—	[2, 4]	—	—
	2	[4, 6]	[3, 5]	—	[1, 3]	—	[2, 3]	—
	3	[4, 6]	[6, 7]	—	—	—	—	[1, 3]
2	1	[2, 4]	[3, 4]	[1, 2]	—	[1, 3]	—	—
	2	2	1	—	[1, 2]	—	[1, 3]	—
	3	[3, 4]	[3, 5]	—	—	—	—	[1, 2]
3	1	[3, 5]	[4, 6]	[1, 2]	—	[1, 2]	—	—
	2	[5, 7]	[3, 6]	—	[1, 2]	—	[1, 2]	—
	3	[3, 4]	[7, 9]	—	—	—	—	[2, 3]
4	1	[4, 7]	[6, 9]	[1, 2]	—	[2, 2]	—	—
	2	[5, 8]	[5, 7]	—	[1, 3]	—	[3, 4]	—
	3	[5, 8]	[3, 5]	—	—	—	—	[1, 2]
5	1	[5, 7]	[6, 8]	[1, 3]	—	[3, 5]	—	—
	2	[5, 7]	[5, 6]	—	[2, 3]	—	[2, 4]	—
	3	[4, 7]	[4, 7]	—	—	—	—	[3, 4]

TABLE 3. Upper bound efficiencies and optimal values of intermediate products for 5 DMUs.

DMU	UPPER BOUND EFF	LOWR BOUND EFF	w^p	t_j^p	γ_j^p	$\hat{z}_j^{p*}$
1	1	0.244	$w^1 = 0.1\text{E-}4$ $w^2 = 0.1\text{E-}4$	$t_4^1 = 0.1\text{E-}4$ $t_1^2 = 0.1\text{E-}4$	$\gamma_4^1 = \gamma_1^2 = 1$ $\gamma_j^p = 0$; for the rest	$\hat{z}_4^1 = \bar{z}_4^1$ $\hat{z}_1^2 = \bar{z}_1^2$ $\hat{z}_j^p = z_j^p$; for the rest
2	1	0.345	$w^1 = 0.1\text{E-}4$ $w^2 = 0.286$	$t_5^2 = 0.286$	$\gamma_5^2 = 1$ $\gamma_j^p = 0$; for the rest	$\hat{z}_5^2 = \bar{z}_5^2$ $\hat{z}_j^p = z_j^p$; for the rest
3	0.929	0.229	$w^1 = 0.179$ $w^2 = 0.1\text{E-}4$	$t_3^1 = 0.107$ $t_4^1 = 0.179$ $t_5^1 = 0.054$ $t_3^2 = \text{E-}4$	$\gamma_4^1 = \gamma_5^2 = \gamma_3^2 = 1$ $\gamma_3^1 = 0.6$ $\gamma_5^1 = 0.3$ $\gamma_j^p = 0$; for the rest	$\hat{z}_4^1 = \bar{z}_4^1$ $\hat{z}_3^2 = \bar{z}_3^2$ $\hat{z}_5^2 = \bar{z}_5^2$ $\hat{z}_3^1 = 0.6\bar{z}_3^1 + 0.4z_3^1$ $\hat{z}_5^1 = 0.3\bar{z}_5^1 + 0.7z_5^1$ $\hat{z}_j^p = z_j^p$; for the rest
4	0.929	0.202	$w^1 = 0.1\text{E-}4$ $w^2 = 0.1\text{E-}4$	$t_4^1 = 0.1\text{E-}4$	$\gamma_4^1 = 1$ $\gamma_j^p = 0$; for the rest	$\hat{z}_4^1 = \bar{z}_4^1$ $\hat{z}_j^p = z_j^p$; for the rest
5	1	0.434	$w^1 = 0.1\text{E-}4$ $w^1 = 0.119$	$t_1^2 = 0.119$ $t_5^2 = 0.119$	$\gamma_1^2 = \gamma_5^2 = 1$ $\gamma_j^p = 0$; for the rest	$\hat{z}_1^2 = \bar{z}_1^2$ $\hat{z}_5^2 = \bar{z}_5^2$ $\hat{z}_j^p = z_j^p$; for the rest

4. REAL APPLICATION

The agricultural sector encompasses the cultivation of crops and the rearing of livestock to fulfill various human necessities, including sustenance, housing, apparel, and more. It is the biggest industry in the world in terms of both workforce and productivity. Given the steady growth in global population, ensuring adequate food

supplies for all nations has become a top priority. The criticality and urgency of monitoring the agricultural industry and assessing its production units lies in the significant impact it has on countries' economies, not only for satisfying domestic demands but also for exports.

Efficiency evaluation is an integral component of agricultural economics. As global population continues to grow and demand for food increases, it is important to maximize production while minimizing costs. Efficiency evaluation enables the identification of areas of improvement in the production process, allowing producers to optimize resources and increase yields. This evaluation also helps policymakers in formulating and implementing agricultural policies that incentivize farmers to adopt best practices and increase productivity [86].

Moreover, efficiency evaluation aids farmers to stay competitive in the global market by reducing production costs and increasing yields. It also leads to more sustainable agricultural practices, which is critical for preserving natural resources, reducing pollution, and mitigating the impacts of climate change [87]. By improving efficiency, farmers can reduce the use of resources such as water, fertilizers, and pesticides, promoting a more sustainable and eco-friendly agriculture industry.

Overall, the importance of efficiency evaluation in agricultural economics is undeniable. It is an essential tool for improving productivity, reducing costs, and promoting sustainability in the agriculture industry [88]. By utilizing efficiency evaluation techniques, farmers, policymakers, and other stakeholders can work together to ensure a more efficient, productive, and sustainable agriculture industry for the future.

Iran's agricultural sector has traditionally played an important role in its economy, but it has faced significant challenges such as drought, water scarcity, and outdated farming practices. In response, the Iranian government has implemented policies aimed at modernizing the sector and improving productivity, including investments in irrigation infrastructure, subsidies for agricultural inputs, and support for research and development. However, the efficiency of these policies in increasing productivity and reducing production costs remains unclear. Despite the efforts made by the government, the agricultural sector still struggles with low productivity and limited access to credit and markets. This has led to an increased reliance on imports, negatively impacting the country's food security. Therefore, evaluating the efficiency of the production process in the agricultural sector in Iran is crucial for addressing the challenges it faces and achieving sustainable growth in the sector.

Iran possesses extensive agricultural land, of which almost one-third is suitable for farming. However, due to insufficient water distribution and poor soil quality, only a small portion of this land is currently being utilized for agriculture. Rainfed and irrigated farming practices are implemented in different parts of the cultivated area, with rainfed agriculture being limited to areas with sufficient natural water resources and precipitation. Notably, extensive precipitation is only found in the Caspian lowlands, such as Guilan and Mazandaran, making them ideal for rainfed agriculture. Azerbaijan, Khorasan, and Fars also receive adequate winter rains for growing grains, which eliminates the need for additional irrigation. However, the arid plateaus of Iran, as well as its eastern and southeastern regions like Yazd and Hormozgan, are not suitable for rainfed farming and require irrigation for crop growth.

Wheat is a fundamental agricultural commodity in Iran, serving as a staple food for Iranians and playing a significant role in their livelihoods, nutrition, and employment. Iran has a per capita consumption of 160 kg of bread wheat, which exceeds that of most other countries. In order to increase wheat production and meet the population's demand for this essential crop, Iran's government encourages farmers to grow more wheat.

This section aims to analyze the efficiency of wheat farming in Iran's provinces during the 2018–2019 crop year, which lasted from 22 September 2018 to 22 September 2019. Iran had 30 provinces during this period, and to obtain a more precise evaluation of wheat production performance, we used Models (3) and (4) on the network structure shown in Figure 3. The provinces were treated as four-stage network systems, with two parallel processes, each consisting of two series stages. The agricultural operations were divided into two parallel processes, namely irrigation and rainfed, and each parallel process included sowing-growing and harvesting series stages. The data used for this analysis was obtained from the Iranian Ministry of Agricultural Jihad¹, and we illustrated the Iran wheat farming system in Figure 4.

¹www16.maj.ir.

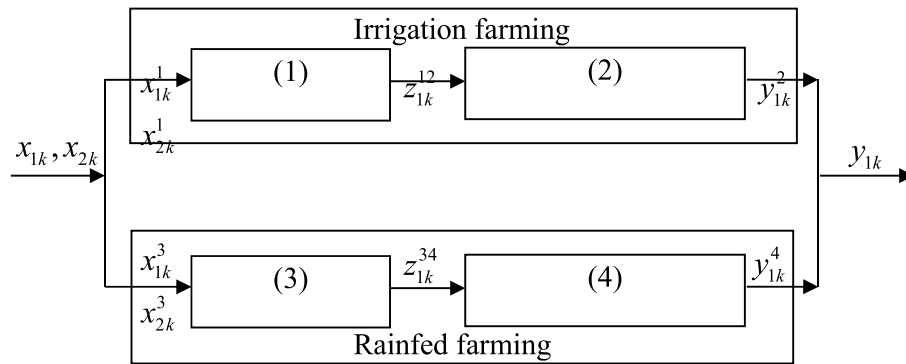


FIGURE 3. Internal structure of each province's agriculture process.

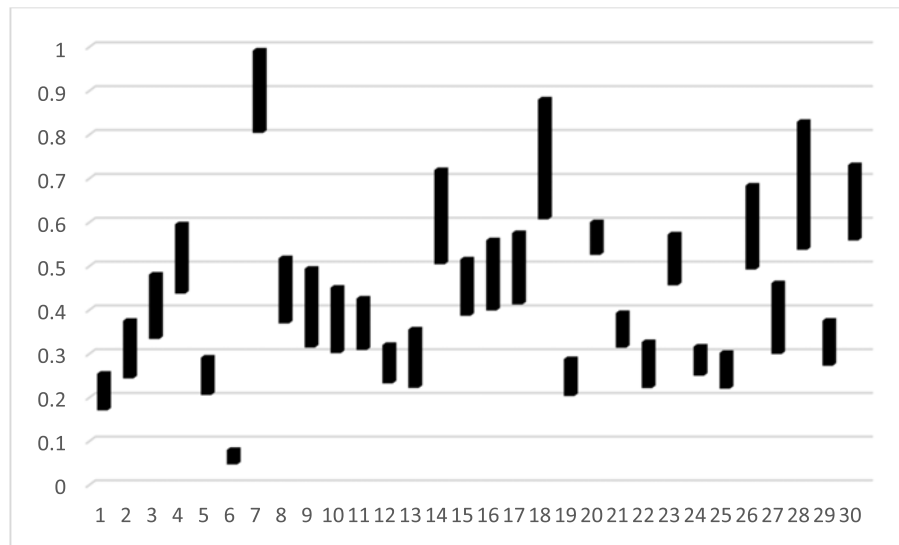


FIGURE 4. Wheat production efficiency in 30 Iranian provinces, 2018–2019.

The complete dataset, which includes interval inputs, intermediate products, and outputs related to wheat farming in the 30 provinces, can be found in Tables 4 and 5. The optimization tasks in this study were carried out using Lingo 18.0 software, with computations performed on a laptop computer equipped with an 11th Gen Intel® Core™ i5 processor running at 2.40 GHz. This configuration guaranteed efficient and dependable performance throughout the analysis. Notably, all process times were negligible.

4.1. Discussion

In the first and third stages of this production analysis, we employ consumed seeds, measured in kilograms, and cultivated area, measured in hectares, as crucial inputs. Harvested area, measured in hectares, plays a pivotal role as an intermediate product, serving as both the output of the first and third stages and the input for the second and fourth stages of the process. Ultimately, the final output from stages 2 and 4 is the production of wheat, measured in tons.

TABLE 4. Upper bound data for 30 provinces.

Provinces	Irrigation farming				Rainfed farming			
	Consumed seed (x_1^1)	Cultivated area (x_2^1)	Harvested area (z_1^{12})	Wheat production (y_1^2)	Consumed seed (x_1^3)	Cultivated area (x_2^3)	Harvested area (z_1^{34})	Wheat production (y_1^4)
1. Azerbaijan, East	110 800	378 608	98 763	270 895	21 875	50 942	354 911	234 319
2. Azerbaijan, West	131 394	385 530	123 065	401 365	24 803	51 248	309 555	350 706
3. Ardabil	98 631	276 371	98 005	481 707	23 241	34 145	206 095	267 550
4. Isfahan	84 440	27 874	70 856	286 416	23 796	2000	26 008	26 003
5. Ilam	47 068	113 874	46 905	151 350	12 555	18 962	69 795	59 214
6. Bushehr	18 914	140 953	18 300	41 861	4000	14 081	16 872	100
7. Tehran	70 755	1400	67 711	361 685	15 539	300	1500	1300
8. Chahar Mahaal and Bakhtiari	31 510	44 236	29 902	117 599	9000	7000	43 317	64 557
9. Khorasan, South	43 598	23 148	43 300	122 740	7600	1000	21 799	8100
10. Khorasan, Razavi	260 456	389 108	257 049	911 702	65 157	21 927	375 015	235 972
11. Khorasan, North	69 032	138 108	67 598	250 003	19 884	19 172	136 979	151 239
12. Khuzestan	512 325	328 864	394 678	1 254 138	128 362	45 393	38 696	27 184
13. Zanjan	29 393	424 267	28 552	119 243	6000	46 216	422 145	451 579
14. Semnan	33 382	8100	32 145	125 625	8000	2500	8000	16 134
15. Sistan and Baluchestan	50 903	200	50 409	126 886	9000	0	0	0
16. Fars	401 673	160 670	360 412	1 407 213	125 603	20 344	118 727	93 436
17. Qazvin	55 976	63 896	55 702	246 499	15 761	6000	63 896	60 050
18. Qom	9000	2000	9000	43 209	2500	350	2000	1900
19. Kurdistan	39 259	526 635	39 259	175 018	10 000	77 125	522 798	516 110
20. Kerman	56 774	300	53 586	181 369	15 022	220	220	250
21. Kermanshah	83 548	358 910	82 461	426 198	17 724	62 601	357 238	417 922
22. Kohgiluyeh and Boyer-Ahmad	28 219	86 326	28 172	81 524	5000	10 500	77 350	71 883
23. Golestan	169 235	231 219	169 023	599 353	35 001	47 642	229 987	522 203
24. Gilan	300	8000	400	600	220	1500	8000	13 328
25. Lorestan	57 437	341 892	53 591	166 182	13 896	54 541	333 503	372 573
26. Mazandaran	14 820	41 869	14 818	55 451	3000	8000	40 843	124 488
27. Markazi	81 944	216 609	80 394	312 969	18 719	23 939	210 885	224 621
28. Hormozgan	16 491	0	15 470	58 020	4200	0	0	0
29. Hamadan	105 464	374 036	100 851	433 373	27 682	51 058	360 696	353 942
30. Yazd	27 267	0	26 902	94 507	5000	0	0	0

Among the provinces under evaluation, it's important to note that Sistan-Baluchestan, Hormozgan, and Yazd do not engage in rainfed farming, and, consequently, their data have been recorded as zero in this context. However, the remaining provinces engage in a combination of both irrigation and rainfed agriculture. Notably, some provinces, such as Azerbaijan-East, Azerbaijan-West, Ardabil, Ilam, Chahar Mahal, Bakhtiari, Golestan, and Khuzestan, exhibit a higher reliance on rainfed farming practices, while others primarily rely on irrigation-based farming methods.

In Table 6, we present interval efficiency data and efficiency classifications for thirty Iranian provinces, focusing primarily on their performance in wheat production. Our comprehensive analysis utilizes the Network DEA models we propose, taking into account the inherent uncertainty associated with interval data. In this table, the upper and lower bounds represent the maximum and minimum efficiency scores achieved by each province, respectively. These bounds are essential in our evaluation as they provide vital efficiency intervals, forming the basis for a holistic ranking of the production units.

Additionally, in Figure 4, we visually depict the efficiency scores of these provinces for the 2018–2019 crop year, using hollow and solid columns to represent the lower and upper efficiency scores, respectively. The chart's horizontal axis displays the provinces, offering a clear view of their respective efficiency scores. This graphical

TABLE 5. Lower bound data for 30 provinces.

Provinces	Irrigation farming				Rainfed farming			
	Consumed seed (x_1^1)	Cultivated area (x_2^1)	Harvested area (z_1^{12})	Wheat production (y_1^2)	Consumed seed (x_1^3)	Cultivated area (x_2^3)	Harvested area (z_1^{34})	Wheat production (y_1^4)
1. Azerbaijan, East	100 800	368 608	94 763	260 895	19 875	47 942	344 911	234 319
2. Azerbaijan, West	121 394	315 530	113 065	391 365	22 803	48 248	299 555	350 706
3. Ardabil	94 631	266 371	94 005	471 707	21 241	32 145	196 095	267 550
4. Isfahan	80 440	25 874	67 856	276 416	21 796	2432	24 008	26 003
5. Ilam	44 068	103 874	43 905	141 350	10 555	16 962	66 795	59 214
6. Bushehr	16 914	130 953	16 300	39 861	3549	12 081	14 872	100
7. Tehran	65 755	1128	65 711	351 685	13 539	82	1087	1300
8. Chahar Mahaal and Bakhtiari	29 510	42 236	27 902	107 599	8556	6508	41 317	64 557
9. Khorasan, South	40 598	21 148	40 300	112 740	7291	482	19 799	8100
10. Khorasan, Razavi	250 456	379 108	247 049	901 702	62 157	19 927	365 015	235 972
11. Khorasan, North	66 032	128 108	65 598	240 003	17 884	17 172	126 979	151 239
12. Khuzestan	502 325	318 864	384 678	1 154 138	118 362	42 393	36 696	27 184
13. Zanjan	27 393	414 267	26 552	109 243	5671	43 216	412 145	451 579
14. Semnan	30 382	7868	30 145	115 625	7414	2086	7868	16 134
15. Sistan and Baluchestan	47 903	10	47 409	116 886	8539	0	0	0
16. Fars	391 673	150 670	350 412	1 307 213	115 603	18 344	108 727	93 436
17. Qazvin	52 976	60 896	52 702	236 499	13 761	5313	60 896	60 050
18. Qom	8399	1707	8387	40 209	2287	146	1689	1900
19. Kurdistan	37 259	516 635	37 259	165 018	9219	73 125	512 798	516 110
20. Kerman	53 774	39	50 586	171 369	13 022	12	21	250
21. Kermanshah	79 548	348 910	78 461	416 198	15 724	59 601	347 238	417 922
22. Kohgiluyeh and Boyer-Ahmad	26 219	83 326	26 172	77 524	4585	9924	74 350	71 883
23. Golestan	159 235	221 219	159 023	589 353	33 001	44 642	219 987	522 203
24. Gilan	105	7595	100	283	13	996	7556	13 328
25. Lorestan	54 437	331 892	50 591	156 182	11 896	51 541	323 503	372 573
26. Mazandaran	12 820	38 869	12 818	52 451	2532	7337	38 843	124 488
27. Markazi	77 944	206 609	77 394	302 969	16 719	21 939	200 885	224 621
28. Hormozgan	13 491	0	13 470	55 020	3838	0	0	0
29. Hamadan	99 464	364 036	97 851	423 373	25 682	48 058	350 696	353 942
30. Yazd	25 267	0	24 902	90 507	4749	0	0	0

representation not only enhances our understanding of efficiency variations among provinces but also highlights their relative ranking status.

Furthermore, we emphasize the importance of working with interval data in our evaluation. Extracting a single value from these intervals could lead to significant data loss. Therefore, our computed efficiency intervals provide a comprehensive ranking for the production units, effectively addressing the inherent uncertainty associated with interval data. Notably, we have included efficiency results generated by the conventional Network DEA Model (1) for the purpose of comparison. These additional findings are now presented in the last column of Table 6, enabling a direct comparison between our proposed model and the conventional approach. It's worth noting that the scores in this column consistently align with the computed efficiency bounds, as expected.

Based on the data presented in Table 6 and Figure 4, Bushehr province consistently demonstrates the lowest efficiency in both lower and upper boundaries, while Tehran province consistently exhibits the highest efficiency in the best cases. These results reinforce our earlier conclusions, consistently identifying Tehran as the efficient province and Bushehr as the fully inefficient.

The largest interval distance is observed for provinces 18 (Qom) and 28 (Hormozgan). The gap between the lower bound and upper bound of efficiency in these two provinces is approximately 0.3. This suggests significant variation in the efficiency scores of these provinces and highlights potential areas for improvement.

TABLE 6. Efficiency interval and classification of 30 provinces.

Provinces	Upper bound	Lower bound	Classification	Model (1) efficiency scores
1. Azerbaijan, East	0.255	0.171	E^-	0.214
2. Azerbaijan, West	0.375	0.244	E^-	0.314
3. Ardabil	0.481	0.334	E^-	0.411
4. Isfahan	0.595	0.437	E^-	0.509
5. Ilam	0.291	0.206	E^-	0.246
6. Bushehr	0.080	0.047	E^{--}	0.064
7. Tehran	0.992	0.804	E^+	0.944
8. Chahar Mahaal and Bakhtiari	0.518	0.370	E^-	0.437
9. Khorasan, South	0.494	0.314	E^-	0.437
10. Khorasan, Razavi	0.451	0.302	E^-	0.389
11. Khorasan, North	0.426	0.309	E^-	0.363
12. Khuzestan	0.320	0.233	E^-	0.273
13. Zanzan	0.356	0.222	E^-	0.297
14. Semnan	0.719	0.504	E^-	0.611
15. Sistan and Baluchestan	0.515	0.387	E^-	0.461
16. Fars	0.559	0.399	E^-	0.470
17. Qazvin	0.576	0.413	E^-	0.483
18. Qom	0.880	0.607	E^-	0.737
19. Kurdistan	0.288	0.203	E^-	0.248
20. Kerman	0.600	0.526	E^-	0.570
21. Kermanshah	0.393	0.314	E^-	0.358
22. Kohgiluyeh and Boyer-Ahmad	0.327	0.221	E^-	0.291
23. Golestan	0.572	0.456	E^-	0.524
24. Gilan	0.317	0.250	E^-	0.305
25. Lorestan	0.302	0.22	E^-	0.260
26. Mazandaran	0.684	0.492	E^-	0.594
27. Markazi	0.461	0.300	E^-	0.386
28. Hormozgan	0.829	0.538	E^-	0.670
29. Hamadan	0.376	0.273	E^-	0.324
30. Yazd	0.731	0.559	E^-	0.656

It's noteworthy that provinces 18 and 28 are the only ones with an upper bound of efficiency greater than the maximum lower bound of efficiency among all provinces. This indicates that these two provinces possess the potential to achieve higher levels of efficiency than any other province in the sample.

Overall, Table 6 serves as a valuable tool for policymakers to identify provinces performing well and those in need of improvement. It also facilitates benchmarking provincial performance and the identification of best practices that can be shared among provinces. Based on the data in Table 6 and Figure 4, it's evident that significant variation exists in the efficiency scores of different provinces, and some provinces have greater potential for improvement than others.

The efficiency intervals in Table 6 enable the classification of provinces into four categories: E^{++} , E^+ , E^- , and E^{--} . These categories are determined based on the upper and lower bounds of the efficiency scores for each province.

As the results indicate, there are no provinces in the E^{++} category in the given data. Tehran is the sole province in the E^+ category, signifying its status as the most efficient province in the sample. E^- includes provinces with lower efficiency scores than Tehran. In this case, all provinces except Tehran fall into this category. Lastly, E^{--} includes provinces with the lowest efficiency scores, where the upper bound of the efficiency score is

less than the lower bound of the efficiency score of the least efficient among the other provinces. Consequently, Bushehr falls into this category for the given data.

The incorporation of these additional details enriches the depth and clarity of our results presentation. This comprehensive approach provides valuable insights into the performance of the production units, effectively capturing the subtleties of interval data and enabling meaningful comparisons with the conventional model.

5. CONCLUSIONS

In summary, this research sought to tackle the challenge of performance measurement in general network production units when dealing with interval data. Our study introduced novel models for assessing the upper and lower bounds of efficiency scores and introduced the concept of interval efficiency for each Decision-Making Unit (DMU). Notably, our findings underscored the critical role played by intermediate products in network systems, emphasizing their inclusion as variables in efficiency score calculations.

To illustrate the practicality and applicability of our proposed approach and models, we conducted an empirical application within the Iranian wheat farming industry. This application demonstrated the feasibility of our methodology in real-world scenarios, offering a comprehensive evaluation of the overall efficiency of production units with general network structures. Future investigations may explore the potential for assessing the efficiency of individual sub-units within these networks.

Moreover, further research avenues could delve into the adaptability of our proposed models in different types of network production systems, such as transportation networks or supply chain networks. Additionally, the study could explore the incorporation of alternative data types, such as fuzzy or uncertain data, to enhance the precision of efficiency score calculations.

In conclusion, this research constitutes a valuable contribution to the realm of performance measurement in network production units. By presenting models that account for both intermediate products and interval data, we lay a solid foundation for future studies to build upon. These future endeavors can explore innovative approaches and models that promise more accurate and efficient performance evaluations of diverse network production systems.

Acknowledgements. The authors extend their gratitude to Nelson Maculan, the Editor-in-Chief, and two anonymous journal reviewers for valuable comments and suggestions that improved the paper.

REFERENCES

- [1] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **2** (1987) 429–444.
- [2] W.W. Cooper, K.S. Park and G. Yu, IDEA and AR-IDEA: models for dealing with imprecise data in DEA. *Manage. Sci.* **45** (1999) 597–607.
- [3] S.H. Kim, C.K. Park and K.S. Park, An application of data envelopment analysis in telephone offices evaluation with partial data. *Comput. Oper. Res.* **26** (1999) 59–72.
- [4] W.W. Cooper, K.S. Park and G. Yu, IDEA (imprecise data envelopment analysis) with CMDs (column maximum decision making units). *J. Oper. Res. Soc.* **52** (2001) 176–181.
- [5] D.K. Despotis and Y.G. Smirlis, Data envelopment analysis with imprecise data. *Eur. J. Oper. Res.* **140** (2002) 24–36.
- [6] T. Entani, Y. Maeda and H. Tanaka, Dual models of interval DEA and its extension to interval data. *Eur. J. Oper. Res.* **136** (2002) 32–45.
- [7] Y.K. Lee, K.S. Park and S.H. Kim, Identification of inefficiencies in an additive model based IDEA (imprecise data envelopment analysis). *Comput. Oper. Res.* **29** (2002) 1661–1676.
- [8] J. Zhu, Imprecise data envelopment analysis (IDEA): a review and improvement with an application. *Eur. J. Oper. Res.* **144** (2003) 513–529.
- [9] K.S. Park, Simplification of the transformations and redundancy of assurance regions in IDEA (imprecise DEA). *J. Oper. Res. Soc.* **55** (2004) 1363–1366.
- [10] J. Zhu, Imprecise DEA via standard linear DEA models with a revisit to a Korean mobile telecommunication company. *Oper. Res.* **52** (2004) 323–329.
- [11] G.R. Jahanshahloo, R. Kazemi-Matin and A. Hadi-Vencheh, On FDH efficiency analysis with interval data. *Appl. Math. Comput.* **159** (2004) 47–55.

- [12] G.R. Jahanshahloo, R. Kazemi-Matin and A. Hadi-Vencheh, On return to scale of fully efficient DMUs in data envelopment analysis under interval data. *Appl. Math. Comput.* **154** (2004) 31–40.
- [13] G.R. Jahanshahloo, F. Hosseinzadeh-Lofti and M. Moradi, Sensitivity and stability analysis in DEA with interval data. *Appl. Math. Comput.* **156** (2004) 463–477.
- [14] A. Amirteimoori and S. Kordrostami, Multi-component efficiency measurement with imprecise data. *Appl. Math. Comput.* **162** (2005) 1265–1277.
- [15] Y.M. Wang, R. Greatbanks and J.B. Yang, Interval efficiency assessment using data envelopment analysis. *Fuzzy Sets Syst.* **153** (2005) 347–370.
- [16] M.S. Haghghat and E. Khorram, The maximum and minimum number of efficient units in DEA with interval data. *Appl. Math. Comput.* **163** (2005) 919–930.
- [17] C. Kao, Interval efficiency measures in data envelopment analysis with imprecise data. *Eur. J. Oper. Res.* **174** (2006) 1087–1099.
- [18] Y.G. Smirlis, E.K. Maragos and D.K. Despotis, Data envelopment analysis with missing values: an interval DEA approach. *Appl. Math. Comput.* **177** (2006) 1–10.
- [19] R. Farzipoor-Saen, A decision model for selecting technology suppliers in the presence of nondiscretionary factors. *Appl. Math. Comput.* **181** (2006) 1609–1615.
- [20] R. Farzipoor-Saen, An algorithm for ranking technology suppliers in the presence of nondiscretionary factors. *Appl. Math. Comput.* **181** (2006) 1616–1623.
- [21] R. Farzipoor-Saen, Suppliers selection in the presence of both cardinal and ordinal data. *Eur. J. Oper. Res.* **183** (2007) 741–747.
- [22] R. Kazemi-Matin, G.R. Jahanshahloo and A. Hadi-Vencheh, Inefficiency evaluation with an additive DEA model under imprecise data, an application on IAUK departments. *J. Oper. Res. Soc. Jpn.* **50** (2007) 163–177.
- [23] K.S. Park, Efficiency bounds and efficiency classifications in DEA with imprecise data. *J. Oper. Res. Soc.* **58** (2007) 533–540.
- [24] G.R. Jahanshahloo, F. Hosseinzadeh-Lofti, M. Rostamy-Malkhalifeh and M. Ahadzadeh-Namin, A generalized model for data envelopment analysis with interval data. *Appl. Math. Model.* **33** (2009) 3237–3244.
- [25] R. Farzipoor-Saen, Supplier selection by the pair of nondiscretionary factors-imprecise data envelopment analysis models. *J. Oper. Res. Soc.* **60** (2009) 1575–1582.
- [26] K.S. Park, Duality, efficiency computations and interpretations in imprecise DEA. *Eur. J. Oper. Res.* **200** (2010) 289–296.
- [27] H. Azizi and H. Ganjeh-Ajirlu, Measurement of the worst practice of decisionmaking units in the presence of non-discretionary factors and imprecise data. *Appl. Math. Model.* **35** (2011) 4149–4156.
- [28] R. Farzipoor-Saen, Media selection in the presence of flexible factors and imprecise data. *J. Oper. Res. Soc.* **62** (2011) 1695–1703.
- [29] R. Farzipoor-Saen, International market selection using advanced data envelopment analysis. *IMA J. Manage. Math.* **22** (2011) 371–386.
- [30] C. Kao and S.T. Liu, Efficiencies of two-stage systems with fuzzy data. *Fuzzy Set Syst.* **176** (2011) 20–35.
- [31] A. Emrouznejad, M. Rostamy-Malkhalifeh, A. Hatami-Marbini and M. Tavana, General and multiplicative non-parametric corporate performance models with interval ratio data. *Appl. Math. Model.* **36** (2012) 5506–5514.
- [32] W. Zhu and Z. Zhou, Interval efficiency of two-stage network DEA model with imprecise data. *INFOR* **51** (2013) 142–150.
- [33] Y. Bouzembrak, H. Allaoui, G. Goncalves, H. Bouchriha and M. Baklouti, A possibilistic linear programming model for supply chain network design under uncertainty. *IMA J. Manage. Math.* **24** (2013) 209–229.
- [34] A. Hadi-Vencheh, A. Hatami-Marbini, Z. Ghelej-Beigi and K. Gholami, An inverse optimization model for imprecise data envelopment analysis. *Optimization* **64** (2015) 2441–2454.
- [35] H. Azizi, S. Kordrostami and A. Amirteimoori, Slacks-based measures of efficiency in imprecise data envelopment analysis: an approach based on data envelopment analysis with double frontiers. *Comput. Ind. Eng.* **79** (2015) 42–51.
- [36] F. He, X. Xu, R. Chen and L. Zhu, Interval efficiency improvement in DEA by using ideal points. *Measurement* **87** (2016) 138–145.
- [37] G.H. Shirdel, S. Ramezani-Tarkhorani and Z. Jafari, A method for evaluating the performance of decision making units with imprecise data using common set of weights. *Int. J. Appl. Comput. Math.* **3** (2017) 411–423.
- [38] G. Wei and K. Wang, A comparative study of robust efficiency analysis and data envelopment analysis with imprecise data. *Expert Syst. Appl.* **81** (2017) 28–38.
- [39] M. Toloo, E. Keshavarz and A. Hatami-Marbini, Dual-role factors for imprecise data envelopment analysis. *Omega* **77** (2018) 15–31.
- [40] R. Mo, H. Huang and L. Yang, An interval efficiency measurement in DEA when considering undesirable outputs. *Complexity* **2020** (2020) 1–12.
- [41] S. Ghobadi, Merging decision-making units with interval data. *RAIRO: Oper. Res.* **55** (2021) S1605–S1631.
- [42] R. Färe and S. Grosskopf, Network DEA. *Soc.-Econ. Plann. Sci.* **34** (2000) 35–49.
- [43] C. Kao, Network data envelopment analysis: a review. *Eur. J. Oper. Res.* **239** (2014) 1–16.
- [44] C. Kao, Network Data Envelopment Analysis Foundations and Extensions, 2nd edition. Springer Cham (2023) 427.
- [45] R. Färe, R. Grabowski, S. Grosskopf and S. Kraft, Efficiency of a fixed but allocatable input: a non-parametric approach. *Econ. Lett.* **56** (1997) 187–193.
- [46] W.D. Cook, M. Hababou and H.J.H. Tuenter, Multicomponent efficiency measurement and shared inputs in data envelopment analysis: an application to sales and service performance in bank branches. *J. Prod. Anal.* **14** (2000) 209–224.

- [47] W.D. Cook and M. Hababou, Sales performance measurement in bank branches. *Omega* **29** (2001) 299–307.
- [48] G.R. Jahanshahloo, A. Amirteimoori and S. Kordrostami, Multicomponent performance, progress and regress measurement and shared inputs and outputs in DEA for panel data: an application in commercial bank branches. *Appl. Math. Comput.* **151** (2004) 1–16.
- [49] M.M. Yu and C.K. Fan, Measuring the cost effectiveness of multimode bus transit in the presence of accident risks. *Transp. Plan. Techn.* **129** (2006) 383–407.
- [50] C. Kao, Efficiency measurement for parallel production systems. *Eur. J. Oper. Res.* **196** (2009) 1107–1112.
- [51] C. Kao, Efficiency decomposition for parallel production systems. *J. Oper. Res. Soc.* **63** (2012) 64–71.
- [52] M.D. Kremantzis, P. Beullens, L.S. Kyrgiakos and J. Klein, Measurement and evaluation of multi-function parallel network hierarchical DEA systems. *Soc.-Econ. Plann. Sci.* **84** (2022) 101428.
- [53] M.D. Troutt, P.J. Ambrose and C.K. Chan, Optimal throughput for multistage input–output processes. *Int. J. Oper. Prod. Manage.* **21** (2001) 148–158.
- [54] A. Amirteimoori and S. Kordrostami, DEA-like models for multi-component performance measurement. *Appl. Math. Comput.* **163** (2005) 735–743.
- [55] K.S. Park and K. Park, Measurement of multi period aggregative efficiency. *Eur. J. Oper. Res.* **193** (2009) 567–580.
- [56] M. Tsutsui and M. Goto, A multi-division efficiency evaluation of U.S. electric power companies using a weighted slacks-based measure. *Soc.-Econ. Plann. Sci.* **43** (2009) 201–208.
- [57] C. Kao, Efficiency decomposition in network data envelopment analysis: a relational model. *Eur. J. Oper. Res.* **192** (2009) 949–962.
- [58] Q.L. Wei and T.S. Chang, Optimal system design series-network DEA models. *J. Oper. Res. Soc.* **52** (2011) 1109–1119.
- [59] L. Fang, Optimal budget for system design series network DEA model. *J. Oper. Res. Soc.* **65** (2014) 1781–1787.
- [60] R. Lin and Q. Liu, Directional distance based efficiency decomposition for series system in network data envelopment analysis. *J. Oper. Res. Soc.* **73** (2021) 1873–1888.
- [61] T.R. Sexton and H.F. Lewis, Two-stage DEA: an application to major league baseball. *J. Prod. Anal.* **19** (2003) 227–249.
- [62] C. Kao and S.N. Hwang, Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. *Eur. J. Oper. Res.* **85** (2008) 418–429.
- [63] Y. Chen, W.D. Cook, N. Li and J. Zhu, Additive efficiency decomposition in two-stage DEA. *Eur. J. Oper. Res.* **196** (2009) 1170–1176.
- [64] L. Liang, Z.Q. Li and W.D. Cook, Data envelopment analysis efficiency in two-stage networks with feedback. *IIE. Trans.* **43** (2011) 309–322.
- [65] J.S. Liu and W.M. Lu, Network-based method for ranking of efficient units in two-stage DEA models. *J. Oper. Res. Soc.* **63** (2012) 1153–1164.
- [66] G. Halkos, N. Tzeremes and S. Kourtzidis, Weight assurance region in two-stage additive efficiency decomposition DEA model: an application to school data. *J. Oper. Res. Soc.* **66** (2015) 696–704.
- [67] S. Aviles-Sacoto, W.D. Cook, R. Imanirad and J. Zhu, Two-stage network DEA: when intermediate measures can be treated as outputs from the second stage. *J. Oper. Res. Soc.* **66** (2015) 1868–1877.
- [68] M. Tavana, M.A. Kaviani, D. Di-Caprio and B. Rahpeyma, A two-stage data envelopment analysis model for measuring performance in three-level supply chains. *Measurement* **78** (2016) 322–333.
- [69] R. Azizi and R. Kazemi-Matin, Ranking two-stage production units in data envelopment analysis. *Asia Pac. J. Oper. Res.* **33** (2016) 1–19.
- [70] G. Halkos, N. Tzeremes and S. Kourtzidis, A unified classification of two-stage DEA models. *Surv. Oper. Res. Man. Sci.* **19** (2014) 1–16.
- [71] M.S. Shahbazifar, R. Kazemi Matin, M. Khounsiavash and F. Koushki, Group ranking of two-stage production units in network data envelopment analysis. *RAIRO: Oper. Res.* **55** (2021) 1825–1840.
- [72] H.F. Lewis and T.R. Sexton, Network DEA: efficiency analysis of organizations with complex internal structure. *Comput. Oper. Res.* **31** (2004) 1365–1410.
- [73] C. Kao and S.N. Hwang, Efficiency measurement for network systems: IT impact on firm performance. *Decis. Support Syst.* **48** (2010) 437–446.
- [74] S. Lozano, Scale and cost efficiency analysis of networks of processes. *Expert Syst. Appl.* **38** (2011) 6612–6617.
- [75] R. Kazemi-Matin and R. Azizi, A unified network-DEA model for performance measurement of production systems. *Measurement* **60** (2015) 186–193.
- [76] F. Boloori, M. Afsharian and J. Pourmahmoud, Equivalent multiplier and envelopment DEA models for measuring efficiency under general network structures. *Measurement* **80** (2016) 259–269.
- [77] A. Kalhor and R. Kazemi-Matin, Performance evaluation of general network production processes with undesirable outputs: a DEA approach. *RAIRO: Oper. Res.* **52** (2018) 17–34.
- [78] F. Yang, Y. Sun, D. Wang and S. Ang, Network data envelopment analysis with two-level maximin strategy. *RAIRO: Oper. Res.* **56** (2022) 2543–2556.
- [79] L.S. Kyrgiakos, G. Klefodimos, G. Vlontzos and P.M. Pardalos, A systematic literature review of data envelopment analysis implementation in agriculture under the prism of sustainability. *Oper. Res. Int. J.* **23** (2023) 7.
- [80] V.H.L. Saputri, W. Sutopo, M. Hisjam and A. Ma'aram, Sustainable agri-food supply chain performance measurement model for GMO and Non-GMO using data envelopment analysis method. *Appl. Sci.* **9** (2019) 1199.

- [81] A. Kord, A. Payan and S. Saati. Network DEA models with stochastic data to assess the sustainability performance of agricultural practices: an application for Sistan and Baluchestan Province in Iran. *J. Math.* **2022** (2022) 1–19.
- [82] L.C. Lu, S.Y. Chiu, Y. Ho and T.H. Chang, Three-stage circular efficiency evaluation of agricultural food production, food consumption, and food waste recycling in EU countries. *J. Clean. Prod.* **343** (2022) 130870.
- [83] A. Kord, A. Payan and S. Saati, Sustainability and optimal allocation of human resource of agricultural practices in Sistan and Baluchestan Province based on network DEA. *J. Math. Extension* **15** (2021) 1–44.
- [84] A. Abbas, C. Zhao, M. Waseem and A.K.R. Ahmad Khan, Analysis of energy input-output of farms and assessment of greenhouse gas emissions: a case study of cotton growers. *Front. Environ. Sci.* **9** (2022) 826838.
- [85] Y. Yang, Q. Zhuang, G. Tian and S. Wei, A management and environmental performance evaluation of China's family farms using an ultimate comprehensive cross-efficiency model (UCCE). *Sustainability* **11** (2019) 6.
- [86] A. Nandy, P.K. Singh and A.K. Singh, Systematic review and meta-regression analysis of technical efficiency of agricultural production systems. *Global Bus. Rev.* **22** (2021) 396–421.
- [87] X. Li, X. Li and J. Jiang, Deep intelligence-driven efficient forecasting for the agriculture economy of computational social systems. *Comput. Intell. Neurosci.* **2022** (2022). DOI: [10.1155/2022/2854675](https://doi.org/10.1155/2022/2854675).
- [88] K. Kočiřová, Application of the DEA on the measurement of efficiency in the EU countries. *Agric. Econ.* **61** (2015) 51–62.



Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at <https://edpsciences.org/en/subscribe-to-open-s2o>.